

INICIAR SESIÓN

NUESTROS PLANES

TODOS LOS
CURSOS

FORMACIONES

CURSOS

PARA
EMPRESAS

ARTÍCULOS DE TECNOLOGÍA > DATA SCIENCE

Conociendo los tipos de aprendizaje de máquina



Yuri Matheus

01/10/2020

Fui contratado por un e-commerce para crear modelos de inteligencia artificial que ayuden a entender mejor a los clientes.

Estos modelos se utilizarán para averiguar cuánto costará el flete de un producto, ver si un cliente comprará o no y agrupar a los clientes con gustos similares.

¿Pero cómo funcionan estos algoritmos? ¿Sólo tenemos uno para hacer todas esas funciones? ¿Tenemos que entrenar cada algoritmo?

El aprendizaje de máquinas

En todos los casos queremos enseñar a la máquina a trabajar con nuestros datos. Por ejemplo, para calcular el precio del flete de un producto. El flete puede ser diferente dependiendo de la tasa de cobro por la distancia y la cantidad de productos.

Esta tasa no tiene un valor fijo, sino que depende de otros factores de mercado como el precio del combustible, los impuestos sobre los productos, entre otras cosas. En cualquier caso, cuando un usuario pide que se calcule el flete, deberíamos ser capaces de estimar un valor.

Sin embargo, ¿cómo podemos estimar el flete si la tasa puede cambiar? Incluso con la tasa cambiando su valor, podemos utilizar los valores de los fletes pasados como una forma de

estimar los fletes futuros. Es decir, podemos tener un modelo, **un algoritmo de inteligencia artificial**, que estimará el flete para nosotros.

Para ello, sólo hay que entrenar este algoritmo con los fletes antiguos, enseñar cómo se calcularon. Cuando se calcula un nuevo flete puede utilizar estos datos para averiguar el valor que queremos.

En este caso, estamos enseñando a la máquina cómo trabajar con nuestros datos, así que estamos supervisando el aprendizaje de la máquina. Entonces, decimos que este algoritmo es un algoritmo de **aprendizaje supervisado**.

En el aprendizaje supervisado pasamos todos los datos para entrenar el modelo. Es decir, pasamos los datos y los resultados. Basándose en esto, el algoritmo encuentra una manera de trabajar con los datos.

Hemos visto que podemos usar este tipo de algoritmos para predecir el precio de los fletes, pero todavía tenemos que clasificar a los clientes y agruparlos por sus gustos, ¿cómo podemos hacerlo?

Clasificando los datos

Tenemos que averiguar si un cliente va a comprar un producto en el sitio o no. Para ello, podemos utilizar alguna información como si viniera de un enlace externo, como un correo electrónico, o un anuncio, si miró la descripción de otro producto, si se ha registrado.

Podemos usar los datos de acceso del usuario para entrenar el modelo para clasificar si el usuario puede comprar o no. Podemos tomar estos datos y hablar con nuestro algoritmo: mira, este usuario vino de un enlace externo, no miró la descripción, se ha registrado y ha comprado el producto.

Pero este otro no vino de un enlace externo, miró la descripción, no tenía registro y compró el producto. Para cada uno de los comportamientos en el sitio, entrenamos la máquina con los datos y con los resultados. Cuando un nuevo usuario accede a la página, podemos decir al modelo cuál fue su comportamiento.

Así, el modelo puede predecir si este nuevo usuario comprará o no. En este caso, como en el del flete, estamos pasando los datos y resultados para entrenar nuestro algoritmo. Sin embargo, a diferencia del otro, no estamos tratando de predecir un valor, sino una **clasificación**.

Buscamos clasificar a un usuario en función de si va a comprar o no. Por lo tanto, este tipo de algoritmo - que también es un algoritmo de aprendizaje supervisado, ya que estamos pasando los datos y resultados para enseñar el modelo - se llama **algoritmo de clasificación**.

Mientras que el algoritmo para calcular el flete se llama de **algoritmo de regresión**. Estos dos tipos de algoritmos son los más conocidos y utilizados, algoritmos para el aprendizaje supervisado.

Todos estos algoritmos, como el aprendizaje de máquina en general, utilizan funciones matemáticas para entrenar los modelos. En nuestro primer ejemplo, podemos utilizar, por ejemplo, la [regresión lineal](#), mientras que para clasificar, podemos usar el [algoritmo de Naive Bayes](#), por ejemplo.

¡Genial! Vimos cómo resolver dos de nuestros problemas. Ahora todo lo que tenemos que hacer es averiguar cómo podemos agrupar los clientes.

Creando grupos

Mirando los algoritmos que conocemos, podemos pensar que la creación de grupos de clientes es un problema de clasificación. De hecho, podemos aplicar la clasificación para crear grupos para nosotros mismos, pero ¿qué grupos tendríamos?

¿Grupos de clientes que compran electrónicos? ¿Grupo de clientes de hasta cuarenta años? ¿Grupo de clientes de la región norte? Podemos usar estos grupos para enviar correos electrónicos especiales a cada uno de ellos, pero ¿qué pasa si estos clientes sólo tienen esto en común?

Es decir, tal vez ni todos los clientes que compran electrónicos o los que viven en una cierta región tienen gustos similares.

Pero entonces, ¿cómo podemos agrupar a los clientes? Hemos visto que podemos clasificar nuestros datos, pero esto podría generar grupos no muy cercanos a lo que queremos. ¿Qué pasa si, en lugar de crear los grupos, le decimos al modelo que agrupe los datos de la manera que mejor le parezca?

Así, a través del algoritmo implementado, el modelo asocia un dato con otro y así crea los grupos.

Al contrario que la clasificación, estamos pasando al modelo sólo los datos y no los resultados. Es decir, no estamos supervisando el aprendizaje, así que decimos que esto es un algoritmo de **aprendizaje no supervisado**.

Diferente del aprendizaje supervisado, en este tipo de aprendizaje nosotros no pasamos los resultados para enseñar el algoritmo. El propio algoritmo en sí mismo toma el conjunto de datos y descubre cómo trabajar con él. Los tipos de algoritmos más utilizados en este tipo de aprendizaje son [clustering](#) y las [reglas de asociación](#). Un ejemplo de algoritmo es [K-Means](#).

Sepa más

Además de este tipo de algoritmos, también existe el llamado aprendizaje por refuerzo. En él la máquina interactúa con el ambiente.

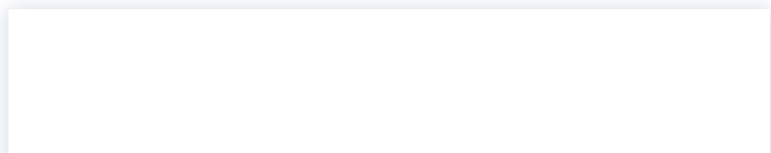
Esta interacción genera una retroalimentación, positiva o negativa. Si es una retroalimentación positiva, la máquina registra que puede volver a realizar esta acción, en caso de retroalimentación negativa, el modelo sabe que no puede volver a realizarla.

Esta fue una breve introducción sobre el tema del aprendizaje de máquina. Para cada conjunto de datos, podemos aplicar diferentes algoritmos y modelos para lograr los mejores resultados.

¿Qué tal aprender más sobre **Machine Learning** y sus diversos recursos? Entonces, ¡Mira nuestros cursos aquí en [Alura](#)!

ARTÍCULOS DE TECNOLOGÍA > DATA SCIENCE

**En Alura encontrarás variados cursos sobre Data Science.
¡Comienza ahora!**



SEMESTRAL

US\$49,90

un solo pago de US\$49,90

- ✓ 218 cursos
- ✓ Videos y actividades 100% en Español
- ✓ Certificado de participación
- ✓ Estudia las 24 horas, los 7 días de la semana
- ✓ Foro y comunidad exclusiva para resolver tus dudas
- ✓ Acceso a todo el contenido de la plataforma por 6 meses

¡QUIERO EMPEZAR A ESTUDIAR!

[Paga en moneda local en los siguientes países](#)

ANUAL

US\$79,90

un solo pago de US\$79,90

- ✓ 218 cursos
- ✓ Videos y actividades 100% en Español
- ✓ Certificado de participación
- ✓ Estudia las 24 horas, los 7 días de la semana
- ✓ Foro y comunidad exclusiva para resolver tus dudas
- ✓ Acceso a todo el contenido de la plataforma por 12 meses

¡QUIERO EMPEZAR A ESTUDIAR!

[Paga en moneda local en los siguientes países](#)

Acceso a todos
los cursos

Estudia las 24 horas,
dónde y cuándo quieras

Nuevos cursos
cada semana

NAVEGACIÓN

PLANES

INSTRUCTORES

BLOG

POLÍTICA DE PRIVACIDAD

TÉRMINOS DE USO

SOBRE NOSOTROS

PREGUNTAS FRECUENTES

¡CONTÁCTANOS!

¡QUIERO ENTRAR EN CONTACTO!

BLOG

PROGRAMACIÓN

FRONT END

DATA SCIENCE

INNOVACIÓN Y GESTIÓN

DEVOPS

AOVS Sistemas de Informática S.A
CNPJ 05.555.382/0001-33

SÍGUENOS EN NUESTRAS REDES SOCIALES



ALIADOS



En Alura somos unas de las Scale-Ups seleccionadas por Endeavor, programa de aceleración de las empresas que más crecen en el país.



Fuimos unas de las 7 startups seleccionadas por Google For Startups en participar del programa Growth Academy en 2021

POWERED BY

CURSOS

Cursos de Programación

Lógica de Programación | Java

Cursos de Front End

HTML y CSS | JavaScript | React

Cursos de Data Science

Data Science | Machine Learning | Excel | Base de Datos | Data Visualization | Estadística

Cursos de DevOps

Docker | Linux

Cursos de Innovación y Gestión

Productividad y Calidad de Vida | Transformación Ágil | Marketing Analytics |
Liderazgo y Gestión de Equipos | Startups y Emprendimiento